BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

REMARKS

Status of The Claims

Claims 46-63, 66-72, 74-78, 80-83, 87-89, 92-95, 98, and 111-122 are pending. Claims 59, 62, 63, 111, 114, and 115 have been allowed. Claims 46-58, 60, 61, 66-72, 74-78, 80-83, 87-89, 92-95, 98, 112, 113, and 116-122 have been rejected.

Claims 53, 80 and 90 were withdrawn from consideration. The Examiner reconsidered the examination of claims 53, 80, and 90, and they are now under examination.

Claims 71, 89, 93 have been amended.

Claims 71(e) and 93 were amended to recite a polypeptide having "up to 147" amino acids. Support for this amendment can be found, for example, in Figure A-2. See, specifically, R6X.

Claim 89 was amended to change dependency.

No new matter has been added by way of these amendments.

The Objections to the Specification Should Be Withdrawn

Page 57 of the specification was objected to for requiring a sequence identifier. The specification has been amended and each primer sequence now has a sequence identifier. A revised sequence listing containing these primers is filed concurrently herewith. The objection has been obviated.

Page 64 of the specification was objected to for having an incomplete citation for Rosenow *et al.* The specification has been amended to remove reference to this paper, and the objection has been obviated.

Page 14, line 15 of the specification recited an amino acid substitution as amino acid "2999." The Examiner is correct that this is a typographical error. The specification has been amended to recite "299," and the objection has been obviated.

The Examiner further requested trademarks appearing in the specification be noted. Pages 57 and 58 have been amended to add the trademark designations.

The Rejection of the Claims Under §112, Second Paragraph, Should Be Withdrawn

Claims 46-58, 60, 66-72, 74-78, 80-83, 87-89, 92-95, 112, and 116-122 were rejected under 35 U.S.C. §112, second paragraph, for indefiniteness. This rejection is respectfully traversed.

I. Claims 46-58, 69-72, 74-78, 80-83, 87-89, 92-94, and 116-122 were rejected for indefiniteness for the recitation of "an amino acid sequence". The Examiner questions if the Applicant intended to claim the entire amino acid sequence or a portion of the sequence. This rejection is respectfully traversed.

The claims at issue recite a polypeptide "comprising an amino acid sequence as set forth in SEQ ID NO:X." This open language is routinely used to claim polypeptides and clearly recites that the polypeptides embraced by the claims will have the amino acid sequence as shown in the recited sequence identifier. Fragments of the recited polypeptides are not encompassed. The claims are clear and the Examiner is respectfully requested to withdraw the rejection.

II. Claims 71 and 72 were rejected for indefiniteness for the recitation of "capable." The Examiner asserts that "an element 'capable' of performing a function is not a positive limitation but only requires the ability to perform." The rejection is respectfully traversed.

Claims 71 and 72 recite, in part, a polypeptide having a recited amino acid sequence "wherein said polypeptide interacts with an antibody, said antibody is capable of interacting with a full-length CbpA polypeptide." The Examiner's attention is drawn to MPEP 2173.05(g) that discusses the acceptability of functional language. The Examiner is respectfully requested to consider *In re Barr*, 170 USPQ 33 (CCPA 1971) which held that the limitation used to define a radical on a chemical compound as "incapable of forming a dye with said oxidizing developing agent" although functional, was perfectly acceptable because it set definite boundaries on the patent protection sought.

In the instant case, the claims at issue employ language that further characterizes the function of an antibody by reciting that the antibody is "capable of interacting with a full-length CbpA polypeptide." To determine the acceptability of claim language, one must determine if

one of skill in the art would understand what is claimed, in view of the specification. The functional language appearing in claims 71 and 72 is clear, and the Examiner is respectfully requested to withdraw the rejection.

III. Claims 51, 52, 60, 66, and 112 were rejected under 35 U.S.C. §112, second paragraph, for indefiniteness. This rejection is respectfully traversed.

The claims were rejected for the term "up to 475 amino acids." Specifically, the Examiner questions if the claim is drawn to a polypeptide that is 0 up to 475 amino acids or 109 amino acids (the length of SEQ ID NO:5) up to 475 amino acids. The confusion appears to stem from the Examiners assertion that the term "an amino acid sequence" as recited in claim 46 can be shorter than the sequence appearing in SEQ ID NO:5. This interpretation of the claim language is incorrect. Claims 46 and 51 are reproduced below.

> 46. An isolated polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:5, wherein said polypeptide does not bind to choline.
>
> 51. The isolated polypeptide of claim 46, wherein said polypeptide comprises an amino acid sequence having up to 475 amino acids.

Claim 46 does not encompass fragments. The claim language "comprising an amino acid sequence as set forth in SEQ ID NO:5" *requires that the polypeptide have the complete amino acid sequence as set forth in SEQ ID NO:5. Additional amino acid sequences can be added at either terminal end.* Claim 51 merely limits the maximum number of amino acids that are added to the terminal ends of the polypeptide of claim 46. An identical line of reasoning applies for claims 52, 60, 66 and 112. Accordingly, the polypeptide recited in claims 51, 52, 60, 66, and 112 is clear, and the Examiner is respectfully requested to withdraw the rejection.

Claims 57, 58, 67, 68, 74-78, 80, 81-83, 87-89, 92-95, 116-122 and claims 70 parts (c), (d), (e), 71 parts (a), (b), (c), (d), (e), and 72 parts (c), (d), (e) were also rejected under 35 U.S.C. §112, second paragraph, for indefiniteness. This rejection is respectfully traversed. As with claims 51, 52, 60, 66, and 112 discussed above, each of these claims also define a length for the recited polypeptide. The reasons outlined above in item III

that clarify the claiming strategy of claims 51, 52, 60, 66, and 112 and are equally applicable to claims 57, 58, 67, 68, 74-78, 80, 81-83, 87-89, 92-95, 116-122 and claims 70 parts (c), (d), (e), 71 parts (a), (b), (c), (d), (e), and 72 parts (c), (d), (e), and the Examiner is respectfully requested to withdraw the rejection.

IV. Claim 89 was also rejected under 35 U.S.C. §112, second paragraph, for indefiniteness. The Examiner states that claim 89 recites a polypeptide having 376 amino acids, but is dependant on claim 87 which recites that the polypeptide has up to 328 amino acids. Claim 89 has been amended to depend from claim 88. As amended, claim 89 is a proper narrowing dependant claim, and the Examiner is respectfully requested to withdraw the rejection.

The Rejection of the Claims Under 35 U.S.C. §102 Should Be Withdrawn

I. Claims 61, 71(f), 98, and 113 were rejected under 35 U.S.C. §102(a) as being anticipated by Hammerschmidt *et al.* (1997) *Molecular Microbiology* 25:1113-1124. The rejection is respectfully traversed.

Claims 61, 71(f) and 113 recite a polypeptide having an amino acid sequence set forth in SEQ ID NO:3, "wherein said amino acid sequence comprises at least one to 57 amino acid substitutions." As illustrated in the Blast alignment provided by the Examiner in the March 30, 2004 Office Action, the sequence disclosed by Hammerschmidt *et al.* comprise more than 57 amino acid substitutions when compared to SEQ ID NO:3. Therefore, Hammerschmidt *et al.* does not anticipate claims 61, 71(f), 98, and 113 and the Examiner is respectfully requested to withdraw the rejection.

II. Claims 54-58, 67-72, 116 and 117 were rejected under 35 U.S.C. §102(e) as being anticipated by U.S. Patent No. 6,573,082. Specifically, the Examiner states that U.S. Patent 6,573,082 discloses SEQ ID NO:38 which corresponds to SEQ ID NO:4 and SEQ ID NO:22 of the instant application. This rejection is respectfully traversed.

Claims 54-58, 67-72, 116 and 117 are drawn to various polypeptides having the amino acid sequence set forth in SEQ ID NO:4 or 22 wherein the "polypeptide does not bind choline." SEQ ID NO:38 disclosed in U.S. Patent 6,573,082 contains a choline-binding domain and therefore does not anticipate the claimed invention. As evidence for the presence of a choline-binding domain and choline-binding activity, Applicants provide in Appendix A Fernandez-Tornero *et al.* (2001) *Nature Structural Biology 8*:1020-1024. Fernandez-Tornero *et al.* studies the structure of choline-binding proteins and provides in Figure 2, page 1021 a description of the consensus sequence for a choline-binding domain. For the Examiner's convenience the Table 1 is provided below which aligns the six (6) choline binding domains appearing in SEQ ID NO:38 with the choline-binding domain consensus sequence determined by Fernandez-Tornero *et al.*

Table 1. Alignment of Choline Binding Domains of SEQ ID NO:38 from U.S. Patent No. 6,573,082 with the Choline Binding Domain Consensus Sequence

Choline binding domain	position	sequence
CBD#1	202-306	G MWYFYNTDGSMATGWLQNNGS WYY
CBC#2	311-326	GAMA.....TGWLQNNGS WYY
CBD#3	331-346	GSMA.....TGWLQNNGS WYY
CBD#4	351-366	GSMA.....TGWLQYNGS WYY
CBD#5	371-386	GSMA.....TGWLQYNGS WYY
CBD#6	391-406	GDMA.....TGWVKD.GDTWYY
Global consensus sequence*		G.m.....TGΦ.....g..WYY

* Φ symbolizes a hydrophobic residue

The evidence presented above clearly establishes that the amino acid sequence set forth in SEQ ID NO:38 of U.S. Patent No. 6,573,082 contains a choline-binding domain and therefore is capable of binding choline. Claims 54-58, 67-72, 116 and 117 recite that the amino acid sequence set forth in SEQ ID NO:4 or 22 "does not bind choline," and accordingly, the claims

are not anticipated by SEQ ID NO:38 of U.S. Patent No. 6,573,082. The Examiner is respectfully requested to withdraw the rejection.

III. Claim 71, part (a), (d), (e) and claims 88 and 93 were rejected under 35 U.S.C. §102(e) as being anticipated by U.S. Patent No. 6,500,613 to Briles *et al.* Specifically, the Examiner asserts U.S. Patent No. 6,500,613 discloses SEQ ID NO:40 and 42 which correspond to SEQ ID NOS:5, 9, and 10 of the present invention. This rejection is respectfully traversed.

Claim 71(d) and claim 88 recite an amino acid sequence set forth in SEQ ID NO:9, "wherein the amino acid sequence comprises at least one to 57 amino acid substitutions ... and comprises up to 376 amino acids." Claims 71(d) and 88 are not anticipated by SEQ ID NO:40 or 42 of U.S. Patent No. 6,500,613.

Claims 71(d) and 88 recite that the polypeptide comprises at least one to 57 amino acid substitutions. Page 13, lines 7-10 of the instant specification define substitutions as "one or more residue specified are replaced by other residues." When SEQ ID NOS:40 and 42 are aligned with SEQ ID NO:9, more than 57 amino acid substitutions occur. For the Examiner's convenience a Clustal W alignment of SEQ ID NO:9 with SEQ ID NO:40 is provided in Appendix B and an alignment of SEQ ID NO:9 with SEQ ID NO:42 is provided in Appendix C. Accordingly, claims 71(d) and 88 are not anticipated by U.S. Patent No. 6,500,613, and the Examiner is respectfully requested to withdraw the rejection.

Claim 71(a) recites an amino acid sequence set forth in SEQ ID NO:5, "wherein the amino acid sequence comprises at least one to 57 amino acid substitutions...and comprises up to 398 amino acids." Claim 71(a) is not anticipated by SEQ ID NO:40 or 42 of U.S. Patent No. 6,500,613.

Claim 71(a) recites that the polypeptide comprises at least one to 57 amino acid substitutions. Page 13, lines 7-10 of the instant specification define substitutions as "one or more residue specified are replaced by other residues". When SEQ ID NOS:40 and 42 are aligned with SEQ ID NO:5, more than 57 amino acid substitutions occur. For the Examiner's convenience a Clustal W alignment of SEQ ID NO:5 with SEQ ID NO:40 is provided in Appendix D, and an alignment of SEQ ID NO:5 with SEQ ID NO:42 is provided in Appendix E.

Accordingly, claim 71(a) is not anticipated by U.S. Patent No. 6,500,613, and the Examiner is respectfully requested to withdraw the rejection.

Claims 71(e) and 93 recite an amino acid sequence set forth in SEQ ID NO:10, "wherein the amino acid sequence comprises at least one to 57 amino acid substitutions...and comprises up to 328 amino acids." Claims 71(e) and 93 are not anticipated by SEQ ID NO:40 or 42 of U.S. Patent No. 6,500,613.

SEQ ID NO:40 (total length 588 amino acids) of U.S. Patent 6,500,613 is a fragment of SEQ ID NO:42 (total length 864 amino acids). Both SEQ ID NO:40 and 42 differ from SEQ ID NO:10 by nine (9) amino acid substitutions. Claims 71(e) and 93 have been amended to recite an amino acid sequence set forth in SEQ ID NO:10, wherein the amino acid sequence comprises at least one to 57 amino acid substitutions and comprises "up to 147 amino acids." Support for this amendment can be found, for example, in Figure 2A-2. U.S. Patent No. 6,500,613 does not disclose a fragment of SEQ ID NO:42 (encoding pspC) that is between 106 and 147 amino acids in length and having the amino acid sequence set forth in SEQ ID NO:10. Accordingly, as amended, claims 71(e) and 93 are not anticipated by U.S. Patent No., 6,500,613 and the rejection should be withdrawn.

CONCLUSIONS

The Examiner is respectfully requested to withdraw the rejections and allow claims 46-58, 60, 61, 66-72, 74-78,m 80-83, 87-89, 92-95, 98, 112, 113, and 116-122. Early notice to this effect is solicited.

It is not believed that extensions of time or fees for net addition of claims are required, beyond those that may otherwise be provided for in documents accompanying this paper. However, in the event that additional extensions of time are necessary to allow consideration of this paper, such extensions are hereby petitioned under 37 CFR § 1.136(a), and any fee required therefore (including fees for net addition of claims) is hereby authorized to be charged to Deposit Account No. 16-0605.

Respectfully submitted,

Kelly J. Williamson

Kelly J. Williamson
Patent Agent
Registration No. 47,179

Customer No. 29312
ALSTON & BIRD LLP
Bank of America Plaza
101 South Tryon Street, Suite 4000
Charlotte, NC 28280-4000
Tel Raleigh Office (919) 862-2200
Fax Raleigh Office (919) 862-2260

"Express Mail" mailing label number: EL868642279US
Date of Deposit June 29, 2004

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to: Mail Stop Amendment, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450

Pamela Lockley

Pamela Lockley

A novel solenoid fold in the cell wall anchoring domain of the pneumococcal virulence factor LytA

Carlos Fernández-Tornero[1], Rubens López[2], Ernesto García[2], Guillermo Giménez-Gallego[1] and Antonio Romero[1]

[1]Departamento de Estructura y Función de Proteínas and [2]Departamento de Microbiología Molecular, Centro de Investigaciones Biológicas - CSIC, C/ Velázquez 144, Madrid, 28006, Spain.

Published online: 5 November 2001, DOI: 10.1038/nsb724

Choline binding proteins are virulence determinants present in several Gram-positive bacteria. Because anchorage of these proteins to the cell wall through their choline binding domain is essential for bacterial virulence, their release from the cell surface is considered a powerful target for a weapon against these pathogens. The first crystal structure of a choline binding domain, from the toxin-releasing enzyme pneumococcal major autolysin (LytA), reveals a novel solenoid fold consisting exclusively of β-hairpins that stack to form a left-handed superhelix. This unique structure is maintained by choline molecules at the hydrophobic interface of consecutive hairpins and may be present in other choline binding proteins that share high homology to the repeated motif of the domain.

Streptococcus pneumoniae (pneumococcus), the leading bacterial cause of acute respiratory infections, is estimated to result in over 6 million deaths every year worldwide from pneumonia, meningitis or bacteremia, especially among children and the elderly[1]. Moreover, the increasing number of antibiotic-resistant strains[2] and the suboptimal clinical efficacy of the available vaccines[3] hamper control of this pathogen. In view of this situation, substantial attention has focused on virulence-related pneumococcal proteins as potential targets for drug design because they are common to all serotypes. Among these proteins, the fifteen-member family[1] of choline binding proteins (ChBPs) appears as a viable target because they are involved in pathogenic processes, such as adhesion to host cells, nasopharyngeal colonization and bacterial sepsis[4]. Although ChBPs are responsible for a wide range of different functions, they all share a highly conserved choline binding domain (ChBD) through which they attach noncovalently to choline moieties of both teichoic and lipoteichoic acids of the cell surface[5]. This manner of displaying proteins at the cell surface, described also for other Gram-positive bacteria and considered peculiar to them[6], is essential for bacterial virulence[7,8].

The major pneumococcal autolysin (LytA), the first and one of the better characterized ChBPs, catalyzes the cleavage of the *N*-acetylmuramoyl-L-alanine bond of the pneumococcal peptidoglycan backbone[9]. LytA is responsible for cellular autolysis, through which it mediates release of toxic substances — such as the pore-forming toxin pneumolysin and cell wall degradation products — that damage endothelial and epithelial barriers and allow pneumococci to gain access to the bloodstream and disseminate through the body[10]. The C-terminal moiety of LytA (C-LytA), consisting of residues Val 188–Lys 318 with six extra amino acids at the N-terminus added during the cloning procedure, was obtained by protein engineering and purification with choline chloride and has been shown to constitute the ChBD of LytA[11] in the fully active form of the enzyme[12]. The primary sequence of C-LytA is constituted by a tandem of imperfect 20 residue repeats, known either as P-motifs[13] or cell wall binding (CWB) repeats[14]. The presence of this sequence repeat defines the family of cell wall binding 1 proteins (Pfam[14] ID code PF01473). Previous analyses of the primary sequence of the fragment suggest that it may contain either six P-motifs[13] or four CWB repeats[14], depending on the nature of the basic 20-residue consensus pattern.

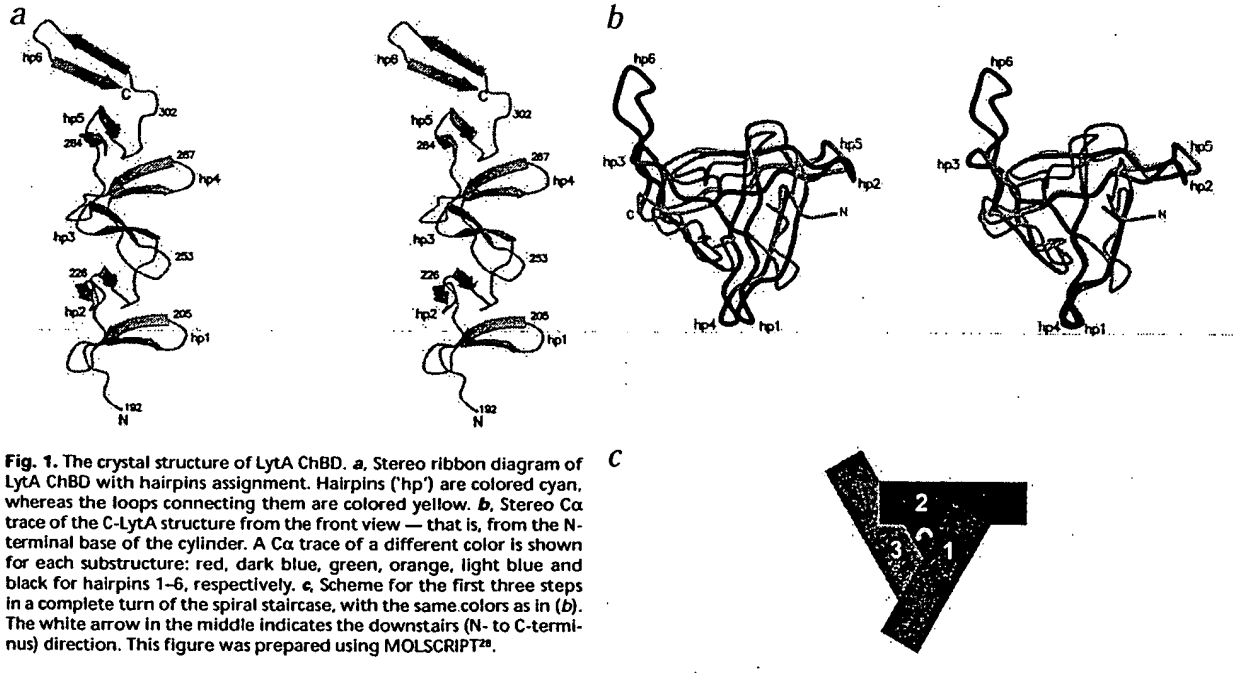

Fig. 1. The crystal structure of LytA ChBD. *a*, Stereo ribbon diagram of LytA ChBD with hairpins assignment. Hairpins ('hp') are colored cyan, whereas the loops connecting them are colored yellow. *b*, Stereo Cα trace of the C-LytA structure from the front view — that is, from the N-terminal base of the cylinder. A Cα trace of a different color is shown for each substructure: red, dark blue, green, orange, light blue and black for hairpins 1–6, respectively. *c*, Scheme for the first three steps in a complete turn of the spiral staircase, with the same colors as in (*b*). The white arrow in the middle indicates the downstairs (N- to C-terminus) direction. This figure was prepared using MOLSCRIPT[28].

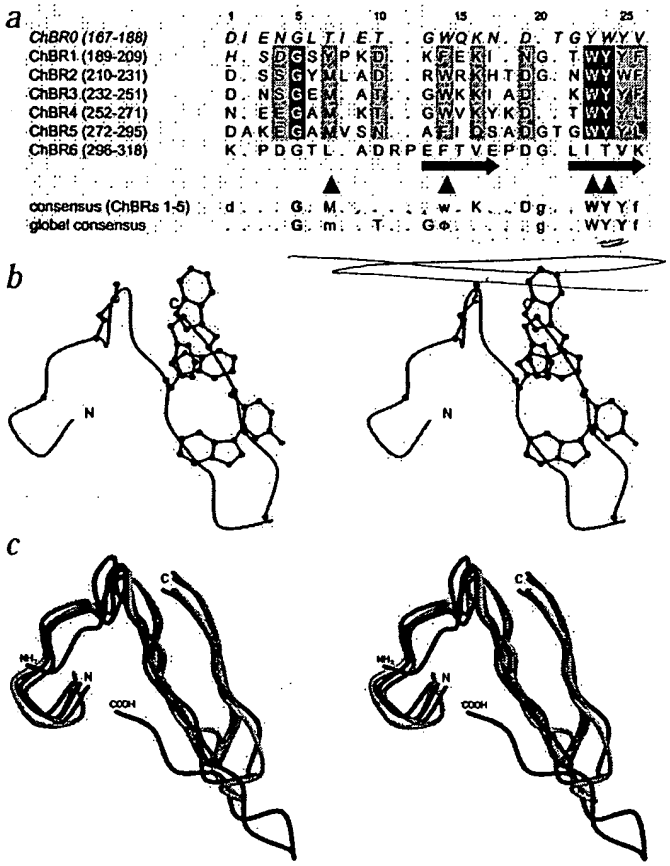

Fig. 2. Sequence and structural similarities among repeats. *a*, Sequence alignment of the seven ChBRs of LytA using the structure criterion and prepared with ALSCRIPT[29]. The repeat numbers and the corresponding ranges of amino acids are shown on the left. The portions of the sequence that form the first and second strands of the hairpins are marked with a barreled arrow below. Conserved residues among ChBRs 1–5 are highlighted yellow (>50% conservation) or red (100% conservation). Choline-binding residues are indicated with a black triangle. A consensus for ChBRs 1–5 has been derived, with bold letters used for 100% conservation, capital letters for 80% and small letters for 60%. The consensus derived from this alignment was used to search for further ChBRs in the primary structure of LytA. The search revealed that the N-terminus of C-LytA may contain a seventh motif (ChBR0). Italicized residues are not visible in the electron density maps (ChBR0 is not even present in the purified protein). The general consensus derived from the >600 CWB repeats found in the Pfam web page[14] has also been included. Φ symbolizes a hydrophobic residue. *b*, Role of the conserved residues (*a*) in the structural motif. ChBR2 has been chosen as example. Only the Cα and side chains of conserved residues common to the two consensus sequences have been represented with ball-and-stick format. *c*, Stereo view of the superposition of the six ChBRs. The same color scheme as in Fig. 1*b* has been used.

In order to show how ChBPs are anchored to the cell wall of Gram-positive bacteria and to assist in the design of new drugs against the infections of these pathogens, we have solved the structure of the C-LytA–choline complex at 2.6 Å resolution using the multiwavelength anomalous dispersion (MAD) method.

Architecture of LytA ChBD

The overall shape of the C-LytA monomer is approximately cylindrical (Fig. 1*a*), with a diameter of 25 Å and height of ~60 Å. The secondary structure is comprised of six independent β-hairpins, labeled by their position in the primary sequence and each consisting of two antiparallel β-strands connected by a short internal loop region (Fig. 1*a*, cyan). Analysis of the secondary structure revealed that all the β-strands have the same length and character (five residues and predominantly hydrophobic). Consecutive hairpins are connected by loops of 8–10 residues (Fig. 1*a*, yellow) that contain a type I +G1 β-bulge turn, plus 4–6 residues mostly in an extended conformation. The hairpins extend perpendicularly from the axis towards the surface of the cylinder, as shown by a frontal view of the protein backbone from the N-terminal base of the cylinder (Fig. 1*b*). With each successive hairpin, a 120° counter-clockwise rotation is introduced so that the i and i+3 hairpins become superimposed, resulting in a left-handed superhelix. Thus, the backbone structure can be described as a spiral staircase with three steps per turn (Fig. 1*c*). The pitch of the i+3 hairpins superhelix is ~30 Å; climbing down the staircase, each hairpin step would lower us ~10 Å.

Based on the described topology, we propose that the ChBD of LytA belongs to the recently defined protein three-dimensional structural family of solenoids, which are structures that contain a superhelical arrangement of repeating structural units[15]. According to Kobe-Kajava classification[15], this new fold could be designated as the left-handed ββ-3-solenoid. As in other known pure β-solenoids[15], no curvature and practically no twist are observed along the staircase axis. Despite these similarities, the ChBD of LytA differs from described β-solenoids because it is built from individual supersecondary bricks — the hairpins — that have their own entity. The currently described structure represents a novel protein fold, as revealed by DALI[16].

Relationship to sequence repeats

The canonical repeat in C-LytA has been proposed to be ~20 residues long, but the repetitive nature of the motifs have made precisely fixing their exact limits by sequence analysis difficult[13,14,17]. Based on the structure presented here, each repeat is proposed to encompass two structural units: a β hairpin and its preceding 8–10-residue connecting loop. Therefore, the six repeats of C-LytA can be redefined structurally as choline binding repeats (ChBRs) 1–6 (Fig. 2*a*). Identifying highly conserved residues based on the structural alignment of the first five repeats is possible, thereby defining a sequence pattern that agrees with the global consensus obtained from a multiple alignment of >600 CWB repeats[14] (Fig. 2*b*). In the second strand of the hairpin, aromatic residues are strictly conserved, whereas less conservation is seen in the N-terminal strand. The preceding loop would be expected to show reduced sequence identity. However, a Gly residue is always found at position 5 of the motif, not only because of topological requirements (a γ_L conformation) of the type I turn but also due to steric hindrance with neighboring side chains.

Although the ChBRs secondary structures are very similar (Fig. 2*c*), deeper analysis of the structure shows that hairpin 6 significantly deviates from the tandem repeats. The angle between hairpin 5 and 6 is only 95° (instead of 120° for all the other hairpins), and the later hairpin is not exactly perpendicular to the cylinder axis. This produces both the imperfect overlap of hairpin 6 with hairpin 3 (Fig. 1*b*) and a slight bend at the C-terminal end of the cylinder that distorts the general fold of the superhelix (Fig. 1*a*). These singular structural features of

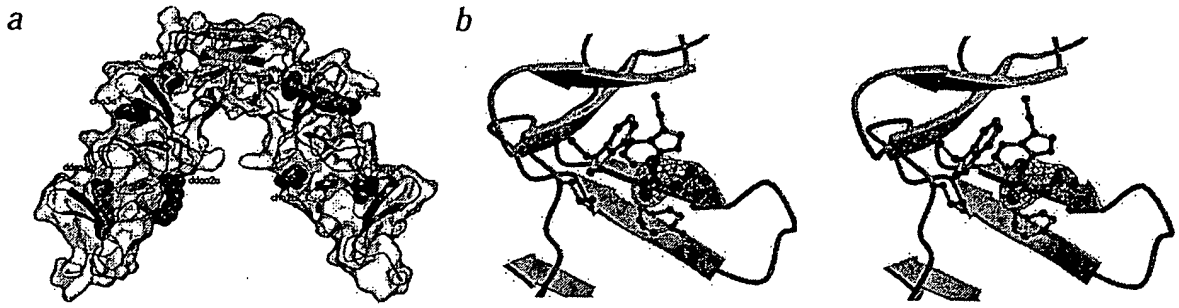

Fig. 3. Choline binding sites. *a*, Ribbon diagram of the C-LytA dimer inscribed into the molecular surface. Monomers are highlighted in different colors: yellow and cyan. ChBSs 1 and 2 of monomer 'a' (yellow) are occupied by DDAO molecules. ChBS3 of monomer 'b' (cyan) is occupied by the (2,2':6',2''-terpyridine)-platinum(II) used for MAD phasing. The hydrophobic components of choline (labeled 'cho'), DDAO (labeled 'ddao') and terpyridin (labeled 'tpy') molecules, schematized as CPK, occupy small hydrophobic cavities on the surface of the protein. *b*, Stereo diagram of ChBS4, where choline is highlighted in orange. The side chains of the hydrophobic conserved residues forming the cavity are shown in the ball-and-stick format. The $2F_o - F_c$ omit map (green) of the choline molecule is contoured at 1.0 σ.

ChBR6 correlate with its unique primary sequence characteristics revealed by alignment of the motifs. The hydrophobic residues in the second strand of its hairpin are less bulky, and this repeat contains a two-residue insertion between the conserved Gly residue of the loop and the beginning of the hairpin (Fig. 2*a*).

Dimer conformation

The asymmetric unit of C-LytA crystals contains two molecules arranged as a dimer throughout their C-terminal regions (Fig. 3*a*). The overall shape of the dimer is reminiscent of a boomerang with arm lengths of 50 Å and an angle between the superhelical arms of ~85°. The arms are related by a noncrystallographic two-fold axis along the bisector of the 85° angle defined by them. The boomerang is likely to carry the catalytic domains of LytA at the end of its arms. The interaction of the monomers involve the predominantly hydrophobic coupling of hairpins 6 with the almost perpendicular pairing of hairpins 5 of both monomers, burying 1,950 Å^2 of surface area per monomer, almost a quarter of the accessible surface area of each monomer (8,600 Å^2). The singular characteristics of the architecture of ChBR6 described in previous sections minimize steric repulsions and introduce the bend, which is necessary to enhance dimerization. Ultracentrifugation has demonstrated that fully active LytA and C-LytA — that is, in the presence of choline chloride — form a dimer in solution, whereas LytA lacking its 16 C-terminal residues forms a monomer in the same conditions[18]. Moreover, the decrease in the catalytic efficiency of LytA (>90%) in this monomeric, truncated form[18] further substantiates the physiological relevance of the C-terminal dimerization.

Choline binding sites

Choline molecules, which form the headgroups of teichoic and lipoteichoic acids in the cell surface, were clearly visible in the electron density maps (Fig. 3*b*). Four choline binding sites (ChBSs) are found per monomer of C-LytA (Fig. 3*a*). Each site is formed by the interface between consecutive hairpin pairs: hairpins 1 and 2 (ChBS1), hairpins 2 and 3 (ChBS2), hairpins 3 and 4 (ChBS3), and hairpins 4 and 5 (ChBS4). The nature of the interaction is mainly hydrophobic, with the three choline methyl groups filling a shallow cavity of ~15 Å^3 constituted by three aromatic residues from the hairpins surrounding the site plus a hydrophobic residue (Met or Leu) from an 8–10 residue connecting loop (Fig. 2*a*, black triangles). A cation–π interaction between the electron-rich systems of the aromatic rings and the positive charge of choline enhances the binding[19]. Other structures of proteins bound to choline or analogs have been determined, including anti phosphoryl-choline[20] and acetylcholine binding proteins[21]. All share similar binding pockets to the hydrophobic head of the choline molecule.

Although the basis of the cation–π interaction implies a common aromatic binding pocket in these structures, choline in the ChBD of LytA appears as an essential requirement for the maintenance of the architecture of the superhelix, shielding the nonburied hydrophobic interface between consecutive hairpins from the solvent and stacking the hairpins together. This idea is supported by experiments showing that the catalytic activity of soluble LytA is only possible upon choline binding[12]. Following the repetitive structure of C-LytA, an additional choline-binding cavity could have been expected to exist between hairpins 5 and 6. Nevertheless, the special architecture of this region (see above) alters the topology of the patch, which then becomes stabilized (especially the residues of hairpin 6) by interaction with the equivalent region of the other monomer upon dimerization. Thus, although choline seems to be required for dimerization[18], our data suggest that it is not directly involved in the assembly of the dimer.

A potential binding surface

Like other biologically relevant helices[15], the structure of the C-LytA monomer displays spiral grooves on its surface that are generated by the left-handed superhelical twist of the molecule (Fig. 4*a*). These grooves, which connect consecutive ChBSs, formed by polar and charged residues (positions 1, 3, 19 and 22 of the alignment; Fig. 2*a*), are ~10 Å long, 7 Å deep and have a maximal aperture of 15 Å. The structure of lipoteichoic acid has been deduced by NMR spectroscopy over three different fragments of the acid and are shown to contain two to eight glycidic building blocks, each of which has two phosphocholine groups[5]. The distance between the hydrophobic heads of these phosphocholines can be reduced to 10 Å by simple torsions. Therefore, the groove between each pair of consecutive ChBSs is likely to accommodate a glycan unit of the pneumococcal teichoic/lipoteichoic acid. Furthermore, N and O atoms from N-acetylated galactosamine residues of the glycan units may establish hydrogen bonds and electrostatic interactions with the corresponding polar atoms of the defined residues on the groove surface (Fig. 4*b*). This could explain the high affinity of C-LytA for pneu-

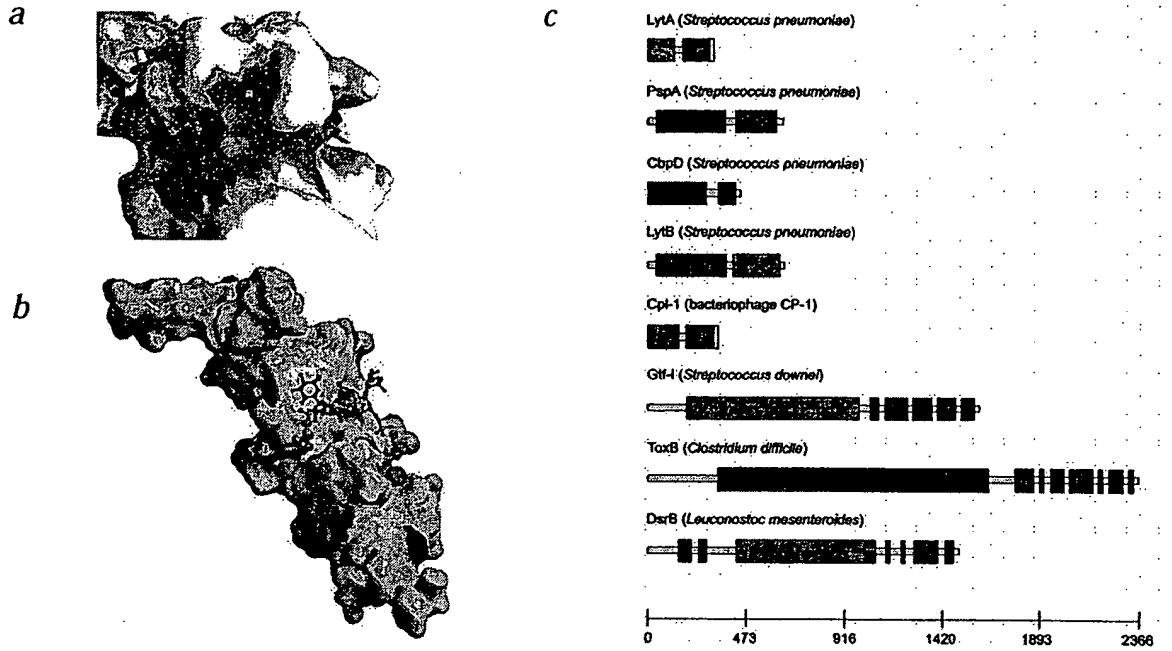

Fig. 4. A potential binding surface and the relevance to other ChBPs. *a*, The groove between two consecutive ChBSs is colored green. Choline molecules are represented with rods. This groove is proposed to shelter the glycan component of teichoic and lipoteichoic acids (*b*). Figure prepared with GRASP[30]. *b*, Scheme of a possible way of anchoring the glycan component of teichoic/lipoteichoic acids in the groove between two consecutive ChBSs. Choline molecules in the currently described structure are schematized as green spheres, whereas a complete repetitive unit of teichoic/lipoteichoic acids is in ball-and-stick representation, with C atoms (gray), N atoms (blue) and O atoms (red). These N and O atoms may establish hydrogen bonds with polar and charged residues (positions 1, 3, 19 and 22 of the alignment; Fig. 2*a*) at the surface of the groove. *c*, Eight illustrative proteins containing ChBRs have been schematized. LytA is the one reported here, PspA has a longer ChBD, CbpD has a lower number of repeats, LytB carries the ChBD at the N-terminus, Cpl-1 is an enzyme of a bacteriophage, Gtf-I of *S. downei* and ToxB of *C. difficile* have a high number of repeats in various tandems, and DsrB of *L. mesenteroides* carries ChBRs at both sides of the catalytic domain. Although the total number of ChBRs (schematized as red-colored boxes) varies greatly among the members of the family (ranging from four to 18), the high sequence conservation of hydrophobic residues present in these repeats (Fig. 2*a*) allows us to propose a common architecture for the ChBD. The catalytic domains with known function are green colored; the blue boxes represent putative functional domains. The last repeats in LytA and Cpl-1 are colored yellow to highlight their particular sequence and structural characteristics.

mococcal cell walls, as the combination of single small binding sites have been shown to provide high affinity to interactions[22].

Relevance to other ChBPs

The primary sequence of the repeating units (not considering ChBR6) is highly conserved within the large number of proteins from Gram-positive bacteria and their bacteriophages (50–100 members, depending on the source) belonging to the cell wall binding 1 family[14]. The total number of repeats and their location in the primary sequence greatly varies among the proteins in which they are present (Fig. 4*c*). Genetic analyses[23] suggest that the ChBD results from a series of gene duplication events that copied the basic repeating unit, as has also been suggested for other solenoid proteins[15]. Because the sequence of the ChBRs is largely defined by the conservation of hydrophobic residues (Fig. 2*a*; alignment of the >600 identified CWB repeats at the Pfam Web page[14]), repeats found in other ChBPs are likely to form a superhelix with the same characteristics as those described here for the ChBD of LytA (Fig. 4*c*). Proteins with a low number of ChBRs in tandem (less than three) are not expected to have affinity for cell walls, because we show that a single choline binding site requires residues from three consecutive repeats. Longer ChBDs are likely to display additional choline binding sites, which would give the protein a higher affinity for the pneumococcal cell wall. In two ChBPs (pneumococcal murein hydrolases LytB and LytC), the ChBD has been found at the N-terminus[14]. There is at least one ChBP (dextransucrase B of *Leuconostoc mesenteroides*) with ChBRs at both sides of the catalytic domain[14]. The solenoid structure described here can easily fit the topology of these three ChBPs, because the overall symmetrical character of the C-LytA monomers with respect to their mass center (ChBR6 not considered) should allow them to accommodate the catalytic domain at either end. The unique sequence and folding characteristics of ChBR6 suggest a divergent evolution. Because ChBR6 is not found in ChBPs other than LytA or those of several pneumococcal bacteriophages (BLAST results not shown), whether the remaining ChBPs dimerize is uncertain.

Conclusions

The first three-dimensional structure for a choline-dependent cell wall anchoring domain, which is present in a wide range of virulence-related proteins from Gram-positive bacteria, provides unique insights into the mechanism of attachment of ChBPs to bacterial cell surface. This structure constitutes a new protein fold, the left-handed ββ-3-solenoid spiral staircase, which consists exclusively of β-hairpins that stack to form a superhelix maintained by choline molecules at hydrophobic cavities on the protein surface. Furthermore, our structure suggests how teichoic and lipoteichoic acids present in the bacterial cell surface may specifically recognize and bind to the ChBD. Because the virulence of pneumococcus is significantly reduced

Table 1 Data collection, phasing and refinement statistics

Data collection	Peak	Inflection	Remote
Wavelength (Å)	1.0695	1.0715	0.9840
Resolution (Å)	35.0–2.6	35.0–2.6	35.0–2.6
Measurements	105,081	96,176	96,859
Unique reflections	11,378	11,382	11,367
R_{sym} (%)[1]	4.9 (26.6)	4.7 (24.8)	5.2 (27.7)
$I / \sigma (I)$[1]	12.3 (2.8)	12.9 (3.0)	8.1 (2.7)
Completeness (%)[1]	99.7 (100)	99.7 (100)	99.6 (100)
MAD phasing			
R_{cullis}			
iso	reference	0.46	0.44
ano	0.74	0.88	0.59
Phasing power			
iso	reference	3.3	3.3
ano	2.2	1.6	2.9
Figure of merit			
Before solvent flattening		0.41	
After solvent flattening		0.84	
Anomalous scatterer		Platinum (1 site)	
Model refinement			
Refinement range (Å)		35–2.6	
Reflections			
Work		10,209	
Free[2]		1,169	
R_{work} / R_{free} (%)		21.8 / 28.2	
R.m.s. deviation			
Bond lengths (Å)		0.007	
Angles (°)		1.2	
Number of atoms			
Protein		2,122	
Solvent		132	
B-factor (Å^2)			
Wilson		43.6	
Average		52.7	

[1]Values in parentheses correspond to the highest resolution shell (2.74–2.60 Å).
[2]Reflexions in the test set represent a 10% of the total number of reflections used during refinement.

when ChBPs are released from the cell wall, the crystal structure of the ChBD from LytA bound to choline may represent a new lead for developing novel drugs against pneumococcal infections. Compounds blocking the ChBSs should emerge as highly effective drugs because they would be aimed toward multiple targets (the entire set of ChBPs), which usually hinders the development of resistances. These drugs are likely to be successful against the Gram-positive bacteria, such as *Clostridium difficile* and *Streptococcus downei*[6], which contain proteins with the ChBD.

Methods

Protein expression, purification and crystallization. The ChBD of the major lytic amidase of *S. pneumoniae* was obtained using established protocols[11]. Crystals were grown using the sitting drop vapor diffusion method at 295 K over a well solution of 30% (w/v) PEG 4000 and 0.2 M ammonium acetate buffered with 0.1 M sodium citrate, pH 6.4, plus 0.4 mM N,N-dimethyl-decylamine-N-oxide (DDAO). These crystals were soaked for 2 h in the same buffer solution containing 4 mM (2,2':6',2"-terpyridine)-platinum(II) chloride.

Data collection and structure determination. Diffraction data were collected at 100 K with a MAR345 detector at DESY-X31 beamline, and processed with MOSFLM[24] and the CCP4 suite[25]. The crystals belong to the I222 space group (a = 58.0 Å, b = 118.2 Å and c = 104.9 Å), with two protein molecules per asymmetric unit and a 56% solvent content (Table1).

The heavy atom search performed with CNS[26] found one platinum site in the asymmetric unit. The Pt-MAD phasing and the subsequent solvent flattening at 2.6 Å were performed with SHARP[27], and the generated electron density map was used to build the first model. Several steps of simulated annealing and B-factor refinement against the peak-dataset using CNS[26] were carried out until the R_{work} and R_{free} values dropped to 21.8% and 28.2%, respectively. The average temperature factor for the N-terminal part of the second monomer (Gly 192–Arg 219), which was modeled with the aid of noncrystallographic symmetry restriction and averaging, is significantly higher (85.21 Å^2) than for the corresponding region of the other monomer (36.86 Å^2).

Coordinates. The coordinates have been deposited in the Brookhaven Protein Data Bank (accession number 1HCX).

Acknowledgments
We thank the staff of beamlines X11, X31 and BW7A, at EMBL-DESY (Hamburg) for support. We are grateful to C. Fernández-Cabrera for excellent technical assistance, P. García and E. Pineda-Molina for helpful discussion, and J.L. García and D. Laurents for critical reading of the manuscript. The first author was supported by a predoctoral fellowship from Ministerio de Educación y Ciencia and by a grant in the Residencia de Estudiantes. This work was partially supported by grants from the Ministerio de Educación y Ciencia of Spain.

Correspondence should be addressed to A.R. *email: romero@cib.csic.es*

Received 26 June, 2001; accepted 21 September, 2001.

1. Tettelin, H. *et al. Science* **293**, 498–506 (2001).
2. Campbell Jr., G.D. & Silberman, R. *Clin. Infect. Dis.* **26**, 1188–1195 (1998).
3. Lipsitch, M. *Emerg. Infect. Dis.* **5**, 336–345 (1999).
4. Hollingshead, S.K. & Briles, D.E. *Curr. Opin. Microbiol.* **4**, 71–77 (2001).
5. Fischer, W. In *Streptococcus pneumoniae* (ed. Tomasz, A.) 155–177 (Mary Ann Liebert, Inc., Larchmont; 2000).
6. Wren, B.W. *Mol. Microbiol.* **5**, 797–803 (1991).
7. Tu, A.T., Fulgham, R.L., McCrory, M.A., Briles, D.E. & Szalai, A.J. *Infect. Immun.* **67**, 4720–4724 (1999).
8. Gosink, K.K., Mann, E.R., Guglielmo, C., Tuomanen, E.I. & Masure, H.R. *Infect. Immun.* **68**, 5690–5695 (2000).
9. Mosser, J.L. & Tomasz, A. *J. Biol. Chem.* **245**, 287–298 (1970).
10. Berry, A.M. & Paton, J.C. *Infect. Immun.* **68**, 133–140 (2000).
11. Sánchez-Puelles, J.M., Sanz, J.M., García, J.L. & García, E. *Gene* **89**, 69–75 (1990).
12. García, E., García, J.L., Ronda, C., García, P. & López, R. *Mol. Gen. Genet.* **201**, 225–230 (1985).
13. García, E. *et al. Proc. Natl. Acad. Sci. USA* **85**, 914–918 (1988).
14. Pfam Web Page *http://www.Sanger.ac.uk/Software/Pfam/* (2001).
15. Kobe, B. & Kajava, A.V. *TIBS* **25**, 509–515 (2000).
16. Holm, L. & Sander, C. *J. Mol. Biol.* **233**, 123 (1993).
17. von Eichel-Streiber, C., Sauerborn, M. & Kuramitsu, H.K. *J. Bacteriol.* **174**, 6707–6710 (1992).
18. Varea, J. *et al. J. Biol. Chem.* **275**, 26842–26855 (2000).
19. Dougherty, D.A. & Stauffer, D.A. *Science* **250**, 1558–1560 (1990).
20. Brown, M., Schumacher, M.A., Wiens, G.D., Brennan, R.G. & Rittenberg, M.B. *J. Exp. Med.* **191**, 2101–2111 (2000).
21. Sussman, J.L. *et al. Science* **253**, 872–979 (1991).
22. Hajduk, P.J., Meadows, R.P. & Fesik, S.W. *Science* **278**, 497–499 (1997).
23. López, R., García, E., García, P. & García, J.L. *Microb. Drug Res.* **3**, 199–211 (1997).
24. Leslie, A.G.W. In *Crystallographic computing V* (eds. Moras, D., Podjarny, A.D. & Thierry, J.C.) 27–38 (Oxford University Press, Oxford; 1991).
25. Collaborative Computational Project, Number 4. *Acta Crystallogr. D* **50**, 760–763 (1994).
26. Brünger, A.T. *et al. Acta Crystallogr. D* **54**, 905–921 (1998).
27. de La Fortelle, E. & Bricogne, G. *Methods Enzymol.* **276**, 472–494 (1997).
28. Kraulis, P. *J. Appl. Crystallogr.* **24**, 946–950 (1996).
29. Barton, G.J. *Protein Eng.* **6**, 37–40 (1993).
30. Nicholls, A., Sharp, K.A. & Honig, B. *Proteins* **11**, 281–296 (1991).

APPENDIX B
SEQ ID NO:9 FROM U.S. App. No. 09/056,019
with SEQ ID NO:40 from U.S. Patent No. 6,500,613

```
CLUSTAL W (1.82) multiple sequence alignment

SEQ TO 40       FASKSERKVHYSIRKFSIGVASVAVASLFLGGVVHAEGVRSGNNLTVTSSGQDISKKYAD 60
SJSEQIDNO9      ------------------------------------------------------------

SEQ             EVESHLESILKDVKKNEKKVAEAQKKVEEAKKKAEDQKEKDRRNYPTITYKTLELEIAES 120
SJSEQIDNO9      ------------------------------------------------------------

SEQ             DVEVKKAELELVKVKAKESQDEEKIKQAEAEVESKQAEATRLKKIKTDREEAKRKADAKL 180
SJSEQIDNO9      ------------------------------------------------------------

SEQ             KEAVEKNVATSEQDKPKRRAKRGVSGELATPDKKENDAKSSDSSVGEETLPSPSLNMAN- 239
SJSEQIDNO9      ----------------------------------ENEG---------STQAATSSNMAKT 17

SEQ             ESQTEHRKDVDEYIKKMLSEIQLDRRKHTQNVNLNIKLSAIKTKYLYELSVLKENSKKEE 299
SJSEQIDNO9      EHRKAAKQVVDEYIEKMLREIQLDRRKHTQNVALNIKLSAIKTKYLRELNVLEEKSK-DE 76

SEQ             LTSKTKAELTAAFEQFKKDTLKPEKKVAEAEKKVEEAKKKAKDQKEEDRRNYPTNTYKTL 359
SJSEQIDNO9      LPSEIKAKLDAAFEKFKKDTLKPGEKVAEAKKKVEEAKKKAEDQKEEDRRNYPTNTYKTL 136

SEQ             ELEIAESDVKVKEAELELVKEEANESRNEEKIKQAKEKVESKKAEATRLEKIKTDRKKAE 419
SJSEQIDNO9      ELEIAEFDVKVKEAELELVKEEAKESRNEGTIKQAKEKVESKKAEATRLENIKTDRKKAE 196

SEQ             EEAKRKAEESEKKAAEAKQKVDAEEYALEAKIAELEYEVQRLEKELKEIDESDSEDYLKE 479
SJSEQIDNO9      EEAKRKADAKLKEANVATS----------------------------DQGKPKGRAKR 226

SEQ             GLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVQLKDAEGNNNVEAYFKEGLEK 539
SJSEQIDNO9      GVPG----------------------------ELATPDKKENDAKSS----------DS 247

SEQ             TTAEKKAELEKAEADLKKAVDEPETPAPAPQPAPAPEKPAEKPAPAPEKPAPAPEKPAPA 599
SJSEQIDNO9      SVGEETL---------------------------------------------------- 254

SEQ             PEKPAPAPEKPAPAPEKPAPTPETPKTGWKQENGMWYFYNTDGSMATGWLQNNGSWYYLN 659
SJSEQIDNO9      ------------------------------------------------------------

SEQ             SNGAMATGWLQNNGSWYYLNSNGAMATGWLQYNGSWYYLNANGDMATGWLQYNGSWYYLN 719
SJSEQIDNO9      ------------------------------------------------------------

SEQ             ANGDMATGWFQYNGSWYYLNANGDMATGWFQYNGSWYYLNANGDMATGWLQYNGSWYYLN 779
SJSEQIDNO9      ------------------------------------------------------------

SEQ             SNGAMVTGWLQNNGSWYYLNANGSMATDWVKDGDTWYYLEASGAMKASQWFKVSDKWYYV 839
SJSEQIDNO9      ------------------------------------------------------------

SEQ             NGSGALAVNTTVDSYRVNANGEWVN 864
SJSEQIDNO9      -------------------------
```

APPENDIX C
SEQ ID NO:9 FROM U.S. App. No. 09/056,019
with SEQ ID NO:42 from U.S. Patent No. 6,500,613

```
CLUSTAL W (1.82) multiple sequence alignment

BRILESSEQIDNO42       EGVRSGNNLTVTSSGQDISKKYADEVESHLESILKDVKKNLKKVQHTQNVGLITKLSEIK 60
SJSEQIDNO9            ------------------------------------------------------------

BRILESSEQIDNO42       KKYLYDLKVNVLSEAELTSKTKETKEKLTATFEQFKKDTLPTEPEKKVAEAQKKVEEAKK 120
SJSEQIDNO9            ------------------------------------------------------------

BRILESSEQIDNO42       KAEDQKEKDRRNYPTITYKTLELEIAESDVEVKKAELELVKVKAKESQDEEKIKQAEAEV 180
SJSEQIDNO9            ------------------------------------------------------------

BRILESSEQIDNO42       ESKQAEATRLKKIKTDREEAKRKADAKLKEAVEKNVATSEQDKPKRRAKRGVSGELATPD 240
SJSEQIDNO9            ------------------------------------------------------------

BRILESSEQIDNO42       KKENDAKSSDSSVGETLPSPSLNMAN-ESQTEHRKDVDEYIKKMLSEIQLDRRKHTQNVN 299
SJSEQIDNO9            --ENE--------GSTQAATSSNMAKTEHRKAAKQVVDEYIEKMLREIQLDRRKHTQNVA 50

BRILESSEQIDNO42       LNIKLSAIKTKYLYELSVLKENSKKEELTSKTKAELTAAFEQFKKDTLKPEKKVAEAEKK 359
SJSEQIDNO9            LNIKLSAIKTKYLRELNVLEEKSK-DELPSEIKAKLDAAFEKFKKDTLKPGEKVAEAKKK 109

BRILESSEQIDNO42       VEEAKKKAKDQKEEDRRNYPTNTYKTLELEIAESDVKVKEAELELVKEEANESRNEEKIK 419
SJSEQIDNO9            VEEAKKKAEDQKEEDRRNYPTNTYKTLELEIAEFDVKVKEAELELVKEEAKESRNEGTIK 169

BRILESSEQIDNO42       QAKEKVESKKAEATRLEKIKTDRKKAEEEAKRKAEESEKKAAEAKQKVDAEEYALEAKIA 479
SJSEQIDNO9            QAKEKVESKKAEATRLENIKTDRKKAEEEAKRKADAKLKEANVATS---------------- 215

BRILESSEQIDNO42       ELEYEVQRLLKELKEIDESDSEDYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAE 539
SJSEQIDNO9            ---------------DQGKPKGRAKRGVPG----------------------------E 231

BRILESSEQIDNO42       IAKLEVQLKDAEGNNNVEAYFKEGLEKTTAEKKAELEKAEADLKKAVDE 588
SJSEQIDNO9            LATPDKKENDAKSS-----------DSSVGEETL--------------- 254
```

APPENDIX D
SEQ ID NO:5 FROM U.S. App. No. 09/056,019
with SEQ ID NO:40 from U.S. Patent No. 6,500,613

```
CLUSTAL W (1.82) multiple sequence alignment

SEQ10 40 40     FASKSERKVHYSIRKFSIGVASVAVASLFLGGVVHAEGVRSGNNLTVTSSGQDISKKYAD 60
SJSEQIDNO5      ----------------------------------------------TEPGE-------- 5
                                                              *..*:

SEQ             EVESHLESILKDVKKNEKKVAEAQKKVEEAKKKAEDQKEKDRRNYPTITYKTLELEIAES 120
SJSEQIDNO5      ------------------KVAEAKKKVEEAEKKAKDQKEEDRRNYPTITYKTLELEIAES 47
                                  *****:******:***:****:**********************

SEQ             DVEVKKAELELVKVKAKESQDEEKIKQAEAEVESKQAEATRLKKIKTDREEAKRKADAKL 180
SJSEQIDNO5      DVEVKKAELELVKVKANEPRDEQKIKQAEAEVESKQAEATRLKKIKTDREEAEEEA---- 103
                ****************:*.:**:*******************************.:*

SEQ             KEAVEKNVATSEQDKPKRRAKRGVSGELATPDKKENDAKSSDSSVGEETLPSPSLNMANE 240
SJSEQIDNO5      ----------------KRRADA---------------------------------------- 109
                                ****.

SEQ             SQTEHRKDVDEYIKKMLSEIQLDRRKHTQNVNLNIKLSAIKTKYLYELSVLKENSKKEEL 300
SJSEQIDNO5      ------------------------------------------------------------

SEQ             TSKTKAELTAAFEQFKKDTLKPEKKVAEAEKKVEEAKKKAKDQKEEDRRNYPTNTYKTLE 360
SJSEQIDNO5      ------------------------------------------------------------

SEQ             LEIAESDVKVKEAELELVKEEANESRNEEKIKQAKEKVESKKAEATRLEKIKTDRKKAEE 420
SJSEQIDNO5      ------------------------------------------------------------

SEQ             EAKRKAEESEKKAAEAKQKVDAEEYALEAKIAELEYEVQRLEKELKEIDESDSEDYLKEG 480
SJSEQIDNO5      ------------------------------------------------------------

SEQ             LRAPLQSKLDTKKAKLSKLEELSDKIDELDAEIAKLEVQLKDAEGNNNVEAYFKEGLEKT 540
SJSEQIDNO5      ------------------------------------------------------------

SEQ             TAEKKAELEKAEADLKKAVDEPETPAPAPQPAPAPEKPAEKPAPAPEKPAPAPEKPAPAP 600
SJSEQIDNO5      ------------------------------------------------------------

SEQ             EKPAPAPEKPAPAPEKPAPTPETPKTGWKQENGMWYFYNTDGSMATGWLQNNGSWYYLNS 660
SJSEQIDNO5      ------------------------------------------------------------

SEQ             NGAMATGWLQNNGSWYYLNSNGAMATGWLQYNGSWYYLNANGDMATGWLQYNGSWYYLNA 720
SJSEQIDNO5      ------------------------------------------------------------

SEQ             NGDMATGWFQYNGSWYYLNANGDMATGWFQYNGSWYYLNANGDMATGWLQYNGSWYYLNS 780
SJSEQIDNO5      ------------------------------------------------------------

SEQ             NGAMVTGWLQNNGSWYYLNANGSMATDWVKDGDTWYYLEASGAMKASQWFKVSDKWYYVN 840
SJSEQIDNO5      ------------------------------------------------------------

SEQ             GSGALAVNTTVDSYRVNANGEWVN 864
```

APPENDIX E
SEQ ID NO:5 FROM U.S. App. No. 09/056,019
with SEQ ID NO:42 from U.S. Patent No. 6,500,613

```
CLUSTAL W (1.82) multiple sequence alignment

BRILESSEQIDNO42     EGVRSGNNLTVTSSGQDISKKYADEVESHLESILKDVKKNLKKVQHTQNVGLITKLSEIK 60
SJSEQIDNO5          ------------------------------------------------------------

BRILESSEQIDNO42     KKYLYDLKVNVLSEAELTSKTKETKEKLTATFEQFKKDTLPTEPEKKVAEAQKKVEEAKK 120
SJSEQIDNO5          ----------------------------------------TEPGEKVAEAKKKVEEAEK 19
                                                            *** :*****:******:*

BRILESSEQIDNO42     KAEDQKEKDRRNYPTITYKTLELEIAESDVEVKKAELELVKVKAKESQDEEKIKQAEAEV 180
SJSEQIDNO5          KAKDQKEEDRRNYPTITYKTLELEIAESDVEVKKAELELVKVKANEPRDEQKIKQAEAEV 79
                    **:****:************************************:*.:**:*********

BRILESSEQIDNO42     ESKQAEATRLKKIKTDREEAKRKADAKLKEAVEKNVATSEQDKPKRRAKRGVSGELATPD 240
SJSEQIDNO5          ESKQAEATRLKKIKTDREEAEEEA--------------------KRRADA---------- 109
                    ********************:.:*                     ****.

BRILESSEQIDNO42     KKENDAKSSDSSVGETLPSPSLNMANESQTEHRKDVDEYIKKMLSEIQLDRRKHTQNVNL 300
SJSEQIDNO5          ------------------------------------------------------------

BRILESSEQIDNO42     NIKLSAIKTKYLYELSVLKENSKKEELTSKTKAELTAAFEQFKKDTLKPEKKVAEAEKKV 360
SJSEQIDNO5          ------------------------------------------------------------

BRILESSEQIDNO42     EEAKKKAKDQKEEDRRNYPTNTYKTLELEIAESDVKVKEAELELVKEEANESRNEEKIKQ 420
SJSEQIDNO5          ------------------------------------------------------------

BRILESSEQIDNO42     AKEKVESKKAEATRLEKIKTDRKKAEEEAKRKAEESEKKAAEAKQKVDAEEYALEAKIAE 480
SJSEQIDNO5          ------------------------------------------------------------

BRILESSEQIDNO42     LEYEVQRLLKELKEIDESDSEDYLKEGLRAPLQSKLDTKKAKLSKLEELSDKIDELDAEI 540
SJSEQIDNO5          ------------------------------------------------------------

BRILESSEQIDNO42     AKLEVQLKDAEGNNNVEAYFKEGLEKTTAEKKAELEKAEADLKKAVDE 588
SJSEQIDNO5          --------------------------------------------------
```